ABC Method applied to brain imaging

Pravin Madhavan

Mathematics and Statistics Centre for Doctoral Training
University of Warwick

December 6th, 2010

Introduction

Clustering of signals is a big topic in neuroimaging.

Clustering of signals is a big topic in neuroimaging.

- Enables us to infer on areas of the brain responsible for given action.

Introduction

Clustering of signals is a big topic in neuroimaging.

- Enables us to infer on areas of the brain responsible for given action.
- Interested in understanding and modelling the dynamics of this process as time increases (discretely).

Introduction

Clustering of signals is a big topic in neuroimaging.

- Enables us to infer on areas of the brain responsible for given action.
- Interested in understanding and modelling the dynamics of this process as time increases (discretely).
- Clustering effect can possibly be modelled by some time dependent statistical model.

Introduction

Two types of statistical models:

Two types of statistical models:

- Prescribed models - likelihood is specified

Introduction

Two types of statistical models:

- Prescribed models - likelihood is specified
- Implicit models - uses some mechanism to simulate observations. No likelihood specified.

Introduction

Two types of statistical models:

- Prescribed models - likelihood is specified
- Implicit models - uses some mechanism to simulate observations. No likelihood specified.
- Implicit models give scientists more freedom to accurately model the phenomenon under consideration.

Introduction

Two types of statistical models:

- Prescribed models - likelihood is specified
- Implicit models - uses some mechanism to simulate observations. No likelihood specified.
- Implicit models give scientists more freedom to accurately model the phenomenon under consideration.
- Interesting results may be found by performing simulations based on model.

Background

Classical Bayesian Monte Carlo Simulation for prescribed models:

Background

Classical Bayesian Monte Carlo Simulation for prescribed models:

- Treat some parameter θ as a random variable, make an informed guess for its prior distribution.

Background

Classical Bayesian Monte Carlo Simulation for prescribed models:

- Treat some parameter θ as a random variable, make an informed guess for its prior distribution.
- Assume likelihood function is known.

Background

Classical Bayesian Monte Carlo Simulation for prescribed models:

- Treat some parameter θ as a random variable, make an informed guess for its prior distribution.
- Assume likelihood function is known.
- Sample from posterior distribution:

$$\pi(\theta \mid D) \propto \text{prior} \times \text{likelihood} = \pi(\theta)\mathbb{P}(D \mid \theta)$$

where $D \subset \mathbb{R}^n$.

Background

Classical Bayesian Monte Carlo Simulation for prescribed models:

- Treat some parameter θ as a random variable, make an informed guess for its prior distribution.
- Assume likelihood function is known.
- Sample from posterior distribution:

$$\pi(\theta \mid D) \propto \text{prior} \times \text{likelihood} = \pi(\theta)\mathbb{P}(D \mid \theta)$$

  where $D \subset \mathbb{R}^n$.
- There are a number of well-understood "classical" Monte Carlo methods for this problem: Rejection sampling, Gibbs sampler, Metropolis-Hastings etc.

Rejection algorithm is given as follows:

Rejection algorithm is given as follows:

- Draw θ from prior $\pi(\cdot)$.

Rejection algorithm is given as follows:

- Draw θ from prior $\pi(\cdot)$.
- Accept θ with probability $\mathbb{P}(D \mid \theta)$.

Drawbacks of classical Monte Carlo methods:

Drawbacks of classical Monte Carlo methods:

- likelihood function may not always be known or may not have a simple form.

Drawbacks of classical Monte Carlo methods:

- likelihood function may not always be known or may not have a simple form.
- Unlikely to find likelihood function for complex systems (namely the clustering problem), so resort to implicit models which require different simulation techniques.

Drawbacks of classical Monte Carlo methods:

- likelihood function may not always be known or may not have a simple form.
- Unlikely to find likelihood function for complex systems (namely the clustering problem), so resort to implicit models which require different simulation techniques.
- This motivates use of ABC method, which does not require any knowledge of likelihood function.

If $\mathbb{P}(D \mid \theta)$ is unknown, perform the ABC Rejection Algorithm given below:

Approximate Bayesian Computation (ABC Method)

If $\mathbb{P}(D \mid \theta)$ is unknown, perform the ABC Rejection Algorithm given below:

- Draw θ from $\pi(\cdot)$.

Approximate Bayesian Computation (ABC Method)

If $\mathbb{P}(D \mid \theta)$ is unknown, perform the ABC Rejection Algorithm given below:

- Draw θ from $\pi(\cdot)$.
- Simulate $D' \sim \mathbb{P}(\cdot \mid \theta)$

Approximate Bayesian Computation (ABC Method)

If $\mathbb{P}(D \mid \theta)$ is unknown, perform the ABC Rejection Algorithm given below:

- Draw θ from $\pi(\cdot)$.
- Simulate $D' \sim \mathbb{P}(\cdot \mid \theta)$
- Accept θ if $D' = D$. The acceptance rate is $P(D)$.

If $\mathbb{P}(D)$ is small, θ will rarely be accepted. A simple modification of the previous algorithm yields a higher acceptance rate.

If $\mathbb{P}(D)$ is small, θ will rarely be accepted. A simple modification of the previous algorithm yields a higher acceptance rate.

- Draw θ from $\pi(\theta)$.

If $\mathbb{P}(D)$ is small, θ will rarely be accepted. A simple modification of the previous algorithm yields a higher acceptance rate.

- Draw θ from $\pi(\theta)$.
- Simulate $D' \sim \mathbb{P}(\cdot \mid \theta)$.

If $\mathbb{P}(D)$ is small, θ will rarely be accepted. A simple modification of the previous algorithm yields a higher acceptance rate.

- Draw θ from $\pi(\theta)$.
- Simulate $D' \sim \mathbb{P}(\cdot \mid \theta)$.
- Accept θ if $d(D, D') \leq \epsilon$.

If $\mathbb{P}(D)$ is small, θ will rarely be accepted. A simple modification of the previous algorithm yields a higher acceptance rate.

- Draw θ from $\pi(\theta)$.
- Simulate $D' \sim \mathbb{P}(\cdot \mid \theta)$.
- Accept θ if $d(D, D') \leq \epsilon$. Note that as $\epsilon \to \infty$, we get observations from the prior $\pi(\theta)$. And if $\epsilon = 0$ we generate observations from $\pi(\theta \mid D)$.

If $\mathbb{P}(D)$ is small, θ will rarely be accepted. A simple modification of the previous algorithm yields a higher acceptance rate.

- Draw θ from $\pi(\theta)$.
- Simulate $D' \sim \mathbb{P}(\cdot \mid \theta)$.
- Accept θ if $d(D, D') \leq \epsilon$. Note that as $\epsilon \to \infty$, we get observations from the prior $\pi(\theta)$. And if $\epsilon = 0$ we generate observations from $\pi(\theta \mid D)$. ϵ reflects the tension between *computability* and *accuracy*.

If data too high dimensional, this is still a far too restrictive acceptance criteria. Reduce dimension by using a summary statistics of the data $S(D)$.

If data too high dimensional, this is still a far too restrictive acceptance criteria. Reduce dimension by using a summary statistics of the data $S(D)$.

- Draw θ from $\pi(\theta)$.

If data too high dimensional, this is still a far too restrictive acceptance criteria. Reduce dimension by using a summary statistics of the data $S(D)$.

- Draw θ from $\pi(\theta)$.
- Simulate $D' \sim \mathbb{P}(\dot{|}\theta)$.

If data too high dimensional, this is still a far too restrictive acceptance criteria. Reduce dimension by using a summary statistics of the data $S(D)$.

- Draw θ from $\pi(\theta)$.
- Simulate $D' \sim \mathbb{P}(\dot{|}\theta)$.
- Accept θ is $d(S(D), S(D')) \leq \epsilon$.

Possible Applications to Clustering Problem

How can these methods be applied to clustering problem?

Possible Applications to Clustering Problem

How can these methods be applied to clustering problem?

- Let θ be array of parameters representing the location of signals (peaks) in 2D snapshot of brain which, under the Bayesian framework, is treated as random variable.

Possible Applications to Clustering Problem

How can these methods be applied to clustering problem?

- Let θ be array of parameters representing the location of signals (peaks) in 2D snapshot of brain which, under the Bayesian framework, is treated as random variable.
- At time step n, derive posterior distribution and sample from it.

Possible Applications to Clustering Problem

How can these methods be applied to clustering problem?

- Let θ be array of parameters representing the location of signals (peaks) in 2D snapshot of brain which, under the Bayesian framework, is treated as random variable.
- At time step n, derive posterior distribution and sample from it.

The posterior distribution at time step n can serve as *prior* at time step $n + 1$.

Possible Applications to Clustering Problem

How can these methods be applied to clustering problem?

- Let θ be array of parameters representing the location of signals (peaks) in 2D snapshot of brain which, under the Bayesian framework, is treated as random variable.
- At time step n, derive posterior distribution and sample from it.

The posterior distribution at time step n can serve as *prior* at time step $n+1$. Hope that clustering effect will take place under reasonable assumptions.